PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 17 March 2000 (17.03.00)	
International application No. PCT/US99/15916	**Applicant's or agent's file reference** BB1179
International filing date (day/month/year) 14 July 1999 (14.07.99)	**Priority date** (day/month/year) 15 July 1998 (15.07.98)
Applicant FALCO, Saverio, Carl et al	

1. The designated Office is hereby notified of its election made:

   [X] in the demand filed with the International Preliminary Examining Authority on:

   07 February 2000 (07.02.00)

   [ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was / [ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland** Facsimile No.: (41-22) 740.14.35	Authorized officer Beate Giffo-Schmitt Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

REC'D 26 OCT 2000
WIPO PCT

15

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference BB1179	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/US99/15916	International filing date *(day/month/year)* 14/07/1999	Priority date *(day/month/year)* 15/07/1998
International Patent Classification (IPC) or national classification and IPC C12N15/53		
Applicant E. I. DU PONT DE NEMOURS AND COMPANY et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 7 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   I ☒ Basis of the report
   II ☐ Priority
   III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   IV ☐ Lack of unity of invention
   V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   VI ☐ Certain documents cited
   VII ☐ Certain defects in the international application
   VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
07/02/2000	25.10.2000
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Armandola, E Telephone No. +49 89 2399 7493

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-24 as originally filed

Claims, No.:

1-12 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application.

☒ claims Nos. 7 (Industrial Applicability).

because:

☒ the said international application, or the said claims Nos. 7 relate to the following subject matter which does not require an international preliminary examination (*specify*):

see separate sheet

☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. are so unclear that no meaningful opinion could be formed (*specify*):

☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.

☐ no international search report has been established for the said claims Nos. .

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	3, 7-12
	No:	Claims	1, 2, 4-6
Inventive step (IS)	Yes:	Claims	7
	No:	Claims	1-6, 8-12
Industrial applicability (IA)	Yes:	Claims	1-6, 8-12
	No:	Claims	

2. Citations and explanations

see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

Re Item III

Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

Industrial Applicability (Art 33 (4) PCT)

Claim 7, in its present formulation, can also refer to a method which could be applied *in vivo* to the human or animal body. As the claim, thus, relates to subject-matter considered by this Authority to be covered by the provisions of Rule 67.1(iv) PCT, no opinion will be formulated with respect to the industrial applicability of the subject-matter of these claims (Art. 34(4)(a)(i) PCT).

For the assessment of the present Claim 7, with regard to *in vivo* methods, on the question whether it is industrially applicable, no unified criteria exist in the PCT Contracting States. The patentability can also be dependent upon the formulation of the claim. The EPO, for example, does not recognize as industrially applicable the subject-matter of claims to the use of a compound in medical treatment, but may allow, however, claims to a known compound for first use in medical treatment and the use of such a compound for the manufacture of a medicament for a new medical treatment.

Re Item V

Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

Reference is made to the following documents:

D1: SASAKI T.: 'AC C26645' EMBL DATABASE,6 August 1997 (1997-08-06), XP002119521 Heidelberg

D2: SASAKI T. ET AL.: 'AC D48925' EMBL DATABASE,9 March 1995 (1995-03-09), XP002119522 Heidelberg

D3: SASAKI T. ET AL.: 'AC D46316' EMBL DATABASE,9 March 1995 (1995-03-09), XP002119523 Heidelberg

D4: PETERSON, CLARENCE DENIS: 'Antifolate studies: the evaluation of folate analogs

as inhibitors of formyltetrahydrofolate synthetase, methenyltetrahydrofolate cyclohydrolase and methylenetetrahydrofolate dehydrogenase' (1979) 109 PP. AVAIL.: UNIV. MICROFILMS INT., ORDER NO. 8000085 FROM: DISS. ABSTR. INT. B 1980, 40(9), 4278, XP002119525

D5: BENNER, M. ET AL: 'Amplification of the methylenetetrahydrofolate reductase--gene' MAIZE GENETICS COOPERATION NEWSLETTER, (1995) NO. 69, PP. 89., [Online] XP002119526 Science and Technology Centre, Rider University, Lawrenceville, New Jersey 08648-3099, USA. Retrieved from the Internet: <URL:www.agron.missouri.edu/mnl/69/101benn er.html> [retrieved on 1999-10-15]

D6: NOUR J M ET AL: 'Isolation and sequencing of the cDNA coding for spinach 10-formyltetrahydrofolate synthetase. Comparisons with the yeast, mammalian, and bacterial proteins.' JOURNAL OF BIOLOGICAL CHEMISTRY, (1992 AUG 15) 267 (23) 16292-6. , XP002119524

D7: WO 95 33054 A (UNIV MCGILL ;ROZEN RIMA (CA); GOYETTE PHILIPPE (CA)) 7 December 1995 (1995-12-07)

D1-D3 disclose nucleic acid sequences, from which the corresponding amino acid sequences have been derived and which are more than 85% identical to the sequences disclosed in the application.

D4 describes the screening of compounds that inhibit tetrathydrofolate metabolism enzymes by incubation of the compounds with purified enzymes.

D5 discloses the amplification of a fragment of maize MTHFR by using degenerate primers derived from the comparison of the sequences of MTHFR of *E. coli, S. typhimurium* and *S. cerevisiae.*

D6 discloses the isolation and sequencing of the cDNA coding for 10-Formyltetrahydrofolate Synthetase from spinach.

D7 describes the isolation of human MTHFR by screening a cDNA library with a probe derived from pig MTHFR.

Novelty (Art. 33(2) PCT)

i) Claims 1, 2 and 4-6 are not novel in view of documents D1-D3 which disclose isolated nucleic acid fragments comprising sequences encoding amino acid stretches more than 85% identical to amino acid sequences of SEQ. ID. NO: 2, 4, 6 and 8. The fragments correspond to functional RNA sequences and have been linked to regulatory sequences and introduced into host cells.
The amino acids coded for by the nucleotide sequences disclosed in D1-D3 comprise a "substantial portion" of a 5, 10 methylentetrahydrofolate reductase peptide with an amino acid sequence selected from SEQ. ID. NO: 2, 4, 6 and 8.

ii) Claims 3 and 7-12 can be considered novel because nucleic acid sequences comprising SEQ. ID. NO: 1, 3, 5, or 7, a method to alter the expression level of a tetrahydrofolate metabolism enzyme as well as methods to obtain a nucleic acid encoding a tetrahydrofolate metabolism enzyme by utilizing a nucleic acid fragment encoding an amino acid sequence of SEQ. ID. NO: 2, 4, 6, or 8 have not been disclosed.

Inventive step (Art 33(3) PCT)

i) Claims 3, 9 and 11 do not entail an inventive step. It was known that plants express enzyme of the tetrahydrofolate metabolism (see D5 and D6). D5 in particular reports the isolation of a fragment of the MTHFR gene from maize. The problem to be solved can, therefore, be seen as the provision of the full length DNA corresponding to the MTHFR gene from maize and/or from other plants and the provision of a method to isolate such DNA.
Document D5 which indicates a method to clone tetrahydrofolate metabolism enzymes from plants utilizing degenerate primers for PCR amplification, obtained from the corresponding enzymes isolated from yeast and bacteria. Besides being clearly indicated in D5, this method is also part of the technology routinely used by the skilled person wanting to isolate related genes in the same or different species.
By using the information found in D5, the skilled person would have arrived at the sequences given in Claim 3 by standard molecular biology methods without the need to exercise inventive skills.

ii) Claims 8 and 10 cannot be considered inventive in view of document D7 which discloses the isolation of hMTHFR by screening a cDNA library with a probe derived from pig MTHFR. The skilled person wanting to isolate further nucleic acids encoding tetrahydrofolate metabolism enzymes would not need to apply inventive ability to apply an analogous method utilizing a different but structurally similar probe.

iii) Claim 12 cannot be considered inventive. D4 describes the screening of folate metabolism enzymes inhibitors in a test utilizing purified enzymes. The difference between the method described in D4 and the subject-matter of Claim 12, is the source of the enzyme(s).
The skilled person wanting to screen antifolate compounds, however, would not have had to exercise inventive activity to apply the method of D4 to any other source of folate metabolism enzymes.

iv) Claim 7 can be considered inventive because none of the available prior art documents suggests the transfection of a gene coding for MTHFR in a host cell for altering the levels of expression an enzyme of the tetrahydrofolate metabolism.

Re Item VIII

Certain observations on the international application

Clarity (Art. 6 PCT)

In claim 6, the relative term " a substantial portion" renders the claim unclear as the meaning of the term "substantial" without any structural or functional limitation is open to individual interpretation and leaves the reader in doubt as to how to identify the portion of the protein covered by the claim.

PATENT COOPERATION TREATY

RECEIVED

NOV 0 1 2000

PATENT RECORDS CENTER

From the INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

TMR

FEULNER, Gregory J.
E.I. DU PONT DE NEMOURS AND COMPANY
Legal/Patent Records Center
1007 Market Street
Wilmington, Delaware 19898
ETATS-UNIS D'AMERIQUE

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Rule 71.1)

Date of mailing *(day/month/year)* 25.10.2000	
Applicant's or agent's file reference BB1179	**IMPORTANT NOTIFICATION**

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US99/15916	14/07/1999	15/07/1998

Applicant
E. I. DU PONT DE NEMOURS AND COMPANY et al.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/	Authorized officer
European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Vullo, C Tel.+49 89 2399-8061

Form PCT/IPEA/416 (July 1992)

15 JAN 2001

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
BB1179		
International application No. PCT/US99/15916	International filing date *(day/month/year)* 14/07/1999	Priority date *(day/month/year)* 15/07/1998
International Patent Classification (IPC) or national classification and IPC C12N15/53		
Applicant E. I. DU PONT DE NEMOURS AND COMPANY et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 7 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
07/02/2000	25.10.2000
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Armandola, E Telephone No. +49 89 2399 7493

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-24 as originally filed

Claims, No.:

1-12 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

☐ the entire international application.

☒ claims Nos. 7 (Industrial Applicability).

because:

☒ the said international application, or the said claims Nos. 7 relate to the following subject matter which does not require an international preliminary examination (*specify*):

see separate sheet

☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. are so unclear that no meaningful opinion could be formed (*specify*):

☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.

☐ no international search report has been established for the said claims Nos. .

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	3, 7-12
	No:	Claims	1, 2, 4-6
Inventive step (IS)	Yes:	Claims	7
	No:	Claims	1-6, 8-12
Industrial applicability (IA)	Yes:	Claims	1-6, 8-12
	No:	Claims	

2. Citations and explanations

see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

Re Item III

Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

Industrial Applicability (Art 33 (4) PCT)

Claim 7, in its present formulation, can also refer to a method which could be applied *in vivo* to the human or animal body. As the claim, thus, relates to subject-matter considered by this Authority to be covered by the provisions of Rule 67.1(iv) PCT, no opinion will be formulated with respect to the industrial applicability of the subject-matter of these claims (Art. 34(4)(a)(i) PCT).

For the assessment of the present Claim 7, with regard to *in vivo* methods, on the question whether it is industrially applicable, no unified criteria exist in the PCT Contracting States. The patentability can also be dependent upon the formulation of the claim. The EPO, for example, does not recognize as industrially applicable the subject-matter of claims to the use of a compound in medical treatment, but may allow, however, claims to a known compound for first use in medical treatment and the use of such a compound for the manufacture of a medicament for a new medical treatment.

Re Item V

Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

Reference is made to the following documents:

D1: SASAKI T.: 'AC C26645' EMBL DATABASE,6 August 1997 (1997-08-06), XP002119521 Heidelberg

D2: SASAKI T. ET AL.: 'AC D48925' EMBL DATABASE,9 March 1995 (1995-03-09), XP002119522 Heidelberg

D3: SASAKI T. ET AL.: 'AC D46316' EMBL DATABASE,9 March 1995 (1995-03-09), XP002119523 Heidelberg

D4: PETERSON, CLARENCE DENIS: 'Antifolate studies: the evaluation of folate analogs

as inhibitors of formyltetrahydrofolate synthetase, methenyltetrahydrofolate cyclohydrolase and methylenetetrahydrofolate dehydrogenase' (1979) 109 PP. AVAIL.: UNIV. MICROFILMS INT., ORDER NO. 8000085 FROM: DISS. ABSTR. INT. B 1980, 40(9), 4278, XP002119525

D5: BENNER, M. ET AL: 'Amplification of the methylenetetrahydrofolate reductase--gene' MAIZE GENETICS COOPERATION NEWSLETTER, (1995) NO. 69, PP. 89., [Online] XP002119526 Science and Technology Centre, Rider University, Lawrenceville, New Jersey 08648-3099, USA. Retrieved from the Internet: <URL:www.agron.missouri.edu/mnl/69/101benn er.html> [retrieved on 1999-10-15]

D6: NOUR J M ET AL: 'Isolation and sequencing of the cDNA coding for spinach 10-formyltetrahydrofolate synthetase. Comparisons with the yeast, mammalian, and bacterial proteins.' JOURNAL OF BIOLOGICAL CHEMISTRY, (1992 AUG 15) 267 (23) 16292-6. , XP002119524

D7: WO 95 33054 A (UNIV MCGILL ;ROZEN RIMA (CA); GOYETTE PHILIPPE (CA)) 7 December 1995 (1995-12-07)

D1-D3 disclose nucleic acid sequences, from which the corresponding amino acid sequences have been derived and which are more than 85% identical to the sequences disclosed in the application.

D4 describes the screening of compounds that inhibit tetrathydrofolate metabolism enzymes by incubation of the compounds with purified enzymes.

D5 discloses the amplification of a fragment of maize MTHFR by using degenerate primers derived from the comparison of the sequences of MTHFR of *E. coli, S. typhimurium* and *S. cerevisiae.*

D6 discloses the isolation and sequencing of the cDNA coding for 10-Formyltetrahydrofolate Synthetase from spinach.

D7 describes the isolation of human MTHFR by screening a cDNA library with a probe derived from pig MTHFR.

Novelty (Art. 33(2) PCT)

i) Claims 1, 2 and 4-6 are not novel in view of documents D1-D3 which disclose isolated nucleic acid fragments comprising sequences encoding amino acid stretches more than 85% identical to amino acid sequences of SEQ. ID. NO: 2, 4, 6 and 8. The fragments correspond to functional RNA sequences and have been linked to regulatory sequences and introduced into host cells.
The amino acids coded for by the nucleotide sequences disclosed in D1-D3 comprise a "substantial portion" of a 5, 10 methylentetrahydrofolate reductase peptide with an amino acid sequence selected from SEQ. ID. NO: 2, 4, 6 and 8.

ii) Claims 3 and 7-12 can be considered novel because nucleic acid sequences comprising SEQ. ID. NO: 1, 3, 5, or 7, a method to alter the expression level of a tetrahydrofolate metabolism enzyme as well as methods to obtain a nucleic acid encoding a tetrahydrofolate metabolism enzyme by utilizing a nucleic acid fragment encoding an amino acid sequence of SEQ. ID. NO: 2, 4, 6, or 8 have not been disclosed.

Inventive step (Art 33(3) PCT)

i) Claims 3, 9 and 11 do not entail an inventive step. It was known that plants express enzyme of the tetrahydrofolate metabolism (see D5 and D6). D5 in particular reports the isolation of a fragment of the MTHFR gene from maize. The problem to be solved can, therefore, be seen as the provision of the full length DNA corresponding to the MTHFR gene from maize and/or from other plants and the provision of a method to isolate such DNA.
Document D5 which indicates a method to clone tetrahydrofolate metabolism enzymes from plants utilizing degenerate primers for PCR amplification, obtained from the corresponding enzymes isolated from yeast and bacteria. Besides being clearly indicated in D5, this method is also part of the technology routinely used by the skilled person wanting to isolate related genes in the same or different species.
By using the information found in D5, the skilled person would have arrived at the sequences given in Claim 3 by standard molecular biology methods without the need to exercise inventive skills.

ii) Claims 8 and 10 cannot be considered inventive in view of document D7 which discloses the isolation of hMTHFR by screening a cDNA library with a probe derived from pig MTHFR. The skilled person wanting to isolate further nucleic acids encoding tetrahydrofolate metabolism enzymes would not need to apply inventive ability to apply an analogous method utilizing a different but structurally similar probe.

iii) Claim 12 cannot be considered inventive. D4 describes the screening of folate metabolism enzymes inhibitors in a test utilizing purified enzymes. The difference between the method described in D4 and the subject-matter of Claim 12, is the source of the enzyme(s).
The skilled person wanting to screen antifolate compounds, however, would not have had to exercise inventive activity to apply the method of D4 to any other source of folate metabolism enzymes.

iv) Claim 7 can be considered inventive because none of the available prior art documents suggests the transfection of a gene coding for MTHFR in a host cell for altering the levels of expression an enzyme of the tetrahydrofolate metabolism.

Re Item VIII
Certain observations on the international application

Clarity (Art. 6 PCT)

In claim 6, the relative term " a substantial portion" renders the claim unclear as the meaning of the term "substantial" without any structural or functional limitation is open to individual interpretation and leaves the reader in doubt as to how to identify the portion of the protein covered by the claim.

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/53, 9/06, 5/10, C12Q 1/68

A1

(11) International Publication Number: WO 00/04163

(43) International Publication Date: 27 January 2000 (27.01.00)

(21) International Application Number: PCT/US99/15916

(22) International Filing Date: 14 July 1999 (14.07.99)

(30) Priority Data:
60/092,869 15 July 1998 (15.07.98) US

(71) Applicant *(for all designated States except US)*: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: FALCO, Saverio, Carl [US/US]; 1902 Millers Road, Arden, DE 19810 (US). FOMODU, Layo, O. [US/US]; 216 Barrett Run Plaza, Newark, DE 19702 (US).

(74) Agent: MAJARIAN, William, R.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: TETRAHYDROFOLATE METABOLISM ENZYMES

(57) Abstract

This invention relates to an isolated nucleic acid fragment encoding a tetrahydrofolate metabolism enzyme. The invention also relates to the construction of a chimeric gene encoding all or a portion of the tetrahydrofolate metabolism enzyme, in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of the tetrahydrofolate metabolism enzyme in a transformed host cell.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/15916

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/53 C12N9/06 C12N5/10 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SASAKI T.: "AC C26645" EMBL DATABASE,6 August 1997 (1997-08-06), XP002119521 Heidelberg the whole document ---	1,2,4-6
X	SASAKI T. ET AL.: "AC D48925" EMBL DATABASE,9 March 1995 (1995-03-09), XP002119522 Heidelberg the whole document ---	1,2,4-6
X	SASAKI T. ET AL.: "AC D46316" EMBL DATABASE,9 March 1995 (1995-03-09), XP002119523 Heidelberg the whole document --- -/--	1,2,4-6

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 October 1999	03/11/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Kania, T

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	NOUR J M ET AL: "Isolation and sequencing of the cDNA coding for spinach 10-formyltetrahydrofolate synthetase. Comparisons with the yeast, mammalian, and bacterial proteins." JOURNAL OF BIOLOGICAL CHEMISTRY, (1992 AUG 15) 267 (23) 16292-6. , XP002119524 the whole document ---	12
Y	PETERSON, CLARENCE DENIS: "Antifolate studies: the evaluation of folate analogs as inhibitors of formyltetrahydrofolate synthetase, methenyltetrahydrofolate cyclohydrolase and methylenetetrahydrofolate dehydrogenase" (1979) 109 PP. AVAIL.: UNIV. MICROFILMS INT., ORDER NO. 8000085 FROM: DISS. ABSTR. INT. B 1980, 40(9), 4278, XP002119525 abstract ---	12
A	BENNER, M. ET AL: "Amplification of the methylenetetrahydrofolate reductase-- gene" MAIZE GENETICS COOPERATION NEWSLETTER, (1995) NO. 69, PP. 89., 'Online! XP002119526 Science and Technology Centre, Rider University, Lawrenceville, New Jerse 08648-3099, USA. Retrieved from the Internet: <URL:www.agron.missouri.edu/mnl/69/101benner.html> 'retrieved on 1999-10-15! abstract ---	1-12
A	WO 95 33054 A (UNIV MCGILL ;ROZEN RIMA (CA); GOYETTE PHILIPPE (CA)) 7 December 1995 (1995-12-07) the whole document ---	1-12
A	COVITZ P. ET AL.: "AC AA660667" EMBL DATABASE, 14 November 1997 (1997-11-14), XP002119527 Heidelberg the whole document -----	1-5

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/15916

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9533054 A	07-12-1995	AU 2519895 A	21-12-1995
		CA 2191220 A	07-12-1995
		EP 0755450 A	29-01-1997

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

TITLE

TETRAHYDROFOLATE METABOLISM ENZYMES

This application claims the benefit of U.S. Provisional Application No. 60/092,869, filed July 15, 1998.

FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding tetrahydrofolate metabolism enzymes in plants and seeds.

BACKGROUND OF THE INVENTION

Tetrahydrofolic acid and its derivatives N^5,N^{10}-methylenetetrahydrofolate, N^5,N^{10}-methenyltetrahydrofolate, N^{10}-formyltetrahydrofolate, and N^5-methyltetrahydrofolate are the biologically active forms of folic acid, oxidized form of tetrahydrofolate (THF). The tetrahydrofolates are coenzymes which are not enzyme-bound and are specialized cosubstrates for a variety of enzymes involved in one-carbon metabolism. Tetrahydrofolate (THF) is a 6-methylpterin derivative linked to p-aminobenzoic acid and glutamic acid residues. Its function is to transfer C1 units in several oxidation states. The C1 units are convalently attached to THF at its N5 and/or N10 positions and enter into the THF pool through the conversion of serine to glycine by serine hydroxymethyl transferase and the cleavage of glycine by glycine synthase. A C1 unit in the THF pool can have several outcomes: it may be used in the conversion of the deoxynucleotide dUMP to dTMP by thymidylate synthase, it may be reduced for the synthesis of methionine, or it may oxidized for the use in the synthesis of purines, since the purines ring of ATP is involved in histidine biosynthesis.

There are several enzymes involved in tetrahydrofolate metabolism five of which are, methylenetetrahydrofolate dehydrogenase (NADP+), 5,10-methylenetetrahydrofolate reductase, 3-methyl-2-oxobutanoate hydroxymethyltransferase, glutamate formyltransferase, or formyltetrahydrofolate deformylase. Methylenetetrahydrofolate dehydrogenase (NADP+) is an oxidoreductase which acts on the CH-NH group of donors with NAD+ or NADP+ as acceptor. In eucaryotes it occurs as a trifunctional enzyme also having methenyltetrahydrofolate cyclohydrolase (EC 3.5.4.9) and formyltetrahydrofolate synthase (EC 6.3.4.3) activity. In some prokaryotes it occurs as a bifunctional enzyme also having methenyltetrahydrofolate cyclohydrolase activity (EC 3.5.4.9). This trifunctional enzyme consists of two major domains: an aminoterminal part, containing the methylene-THF dehydrogenase and methenyl-THF cyclohydrolase activities and a larger formyl-THF synthetase domain.

5,10-Methylenetetrahydrofolate reductase (EC 1.7.99.5) (MTHFR) plays a role in the synthesis of methionine (West et al, (1993) *J. Biol. Chem. 268*:153-160 and D'Ari et al. (1991) *J. Biol. Chem. 266*:23953-23958). S-adenosylmethionine (SAM) an important methyl group donor for many biosynthetic methylation reactions in plants. SAM is formed

from methionine by SAM synthetase. Transfer of the methyl group from SAM to an acceptor molecule results in the formation of S-adenosylhomocysteine, which is then hydrolyzed to homocysteine. Methionine is regenerated from homocysteine by methyl group transfer from 5-methyltetrahydrofolate. This form of folate is generated from 5,10-methylenetetrahydrofolate through the action of 5,10-methylenetetrahydrofolate reductase (MTHFR), a cytosolic flavoprotein. The heavy demand in plant cells for methyl groups derived from SAM necessitate a rapid recycling of S-adenosylhomocysteine, and thus a heavy demand for 5-methyltetrahydrofolate produced by MTHFR.

3-Methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1.2.11) is the first enzyme in the pantothenate biosynthetic pathway. This enzyme catalyses the conversion of 5,10-methylenetetrahydrofolate and 3-methyl-2-oxobutanoate to tetrahydrofolate and 2-dehydropantoate. Pantothenate is a vitamin required in the diet of animals. It is used in the synthesis of coenzyme A, which in turn, is used in many important enzyme reactions in many pathways, e.g., fatty acid biosynthesis. The production of high levels fatty acids, which require coenzyme A for their synthesis, might be stimulated by production of higher levels of coenzyme A, which in turn would require increased production of pantothenate. Another use might be for the increased production of pantothenate in plants in order to purify this vitamin for sale.

Glutamate formyltransferase (EC 2.1.2.5) catalyses the transfer of a formyl group from 5-formyltetrahydrofolate to L-glutamate. This enzyme serves to channel one-carbon units from formiminoglutamate to the folate pool.

Lastly, formyltetrahydrofolate deformylase (EC 3.5.1.10) catalyses the formation of formate and tetrahydrofolate from 10-formyltetrahydrofolate and water. 10-Formyltetrahydrofolate is required in de novo purine biosynthesis and histidine biosynthesis.

Because these enzymes are involved in tetrahydrofolate metabolism, amino acid synthesis, fatty acid biosynthesis and de novo synthesis of purines inhibition of their activity may be lethal, thus suggesting that they would be attractive herbicide targets. Thus production of these plant enzymes in bacteria for use in a high throughput screen for chemical inhibitors would be desirable. Alternatively, overproduction of these enzymes in transgenic plants could be used to enhance the production of many secondary metabolites, amino acids, purine nucleic acids and vitamins. Accordingly, the availability of nucleic acid sequences encoding all or a portion of an enzyme involved in tetrahydrofolate metabolism would facilitate studies to better understand tetrahydrofolate metabolism in plants, provide genetic tools to enhance the production of secondary metabolites, amino acids and vitamins. These enzymes may also provide targets to facilitate design and/or identification of inhibitors tetrahydrofolate metabolism that may be useful as herbicides.

SUMMARY OF THE INVENTION

The instant invention relates to isolated nucleic acid fragments encoding tetrahydrofolate metabolism enzymes. Specifically, this invention concerns an isolated nucleic acid fragment encoding a 5,10-methylenetetrahydrofolate reductase and an isolated nucleic acid fragment that is substantially similar to an isolated nucleic acid fragment encoding a 5,10-methylenetetrahydrofolate reductase. In addition, this invention relates to a nucleic acid fragment that is complementary to the nucleic acid fragment encoding 5,10-methylenetetrahydrofolate reductase. An additional embodiment of the instant invention pertains to a polypeptide encoding all or a substantial portion of a 5,10-methylenetetrahydrofolate reductase.

In another embodiment, the instant invention relates to a chimeric gene encoding a 5,10-methylenetetrahydrofolate reductase, or to a chimeric gene that comprises a nucleic acid fragment that is complementary to a nucleic acid fragment encoding a 5,10-methylenetetrahydrofolate reductase. operably linked to suitable regulatory sequences, wherein expression of the chimeric gene results in production of levels of the encoded protein in a transformed host cell that is altered (i.e., increased or decreased) from the level produced in an untransformed host cell.

In a further embodiment. the instant invention concerns a transformed host cell comprising in its genome a chimeric gene encoding a 5,10-methylenetetrahydrofolate reductase, operably linked to suitable regulatory sequences. Expression of the chimeric gene results in production of altered levels of the encoded protein in the transformed host cell. The transformed host cell can be of eukaryotic or prokaryotic origin, and include cells derived from higher plants and microorganisms. The invention also includes transformed plants that arise from transformed host cells of higher plants. and seeds derived from such transformed plants.

An additional embodiment of the instant invention concerns a method of altering the level of expression of a 5,10-methylenetetrahydrofolate reductase in a transformed host cell comprising: a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a 5,10-methylenetetrahydrofolate reductase; and b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of altered levels of 5,10-methylenetetrahydrofolate reductase in the transformed host cell.

An addition embodiment of the instant invention concerns a method for obtaining a nucleic acid fragment encoding all or a substantial portion of an amino acid sequence encoding a 5,10-methylenetetrahydrofolate reductase.

A further embodiment of the instant invention is a method for evaluating at least one compound for its ability to inhibit the activity of a 5,10-methylenetetrahydrofolate reductase, the method comprising the steps of: (a) transforming a host cell with a chimeric gene

comprising a nucleic acid fragment encoding a 5,10-methylenetetrahydrofolate reductase, operably linked to suitable regulatory sequences; (b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of 5,10-methylenetetrahydrofolate reductase in the transformed host cell; (c) optionally purifying the 5,10-methylenetetrahydrofolate reductase expressed by the transformed host cell; (d) treating the 5,10-methylenetetrahydrofolate reductase with a compound to be tested; and (e) comparing the activity of the 5,10-methylenetetrahydrofolate reductase that has been treated with a test compound to the activity of an untreated 5,10-methylenetetrahydrofolate reductase, thereby selecting compounds with potential for inhibitory activity.

BRIEF DESCRIPTION OF THE SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description and the accompanying Sequence Listing which form a part of this application.

Table 1 lists the polypeptides that are described herein. the designation of the cDNA clones that comprise the nucleic acid fragments encoding polypeptides representing all or a substantial portion of these polypeptides, and the corresponding identifier (SEQ ID NO:) as used in the attached Sequence Listing. The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

TABLE 1

Tetrahydrofolate Metabolism Enzymes

Protein	Clone Designation	SEQ ID NO: (Nucleotide)	SEQ ID NO: (Amino Acid)
5,10-Methylenetetrahydrofolate Reductase	cco1.pk0049.d4	1	2
5,10-Methylenetetrahydrofolate Reductase	rls48.pk0005.b11	3	4
5,10-Methylenetetrahydrofolate Reductase	sfl1.pk0017.d12	5	6
5,10-Methylenetetrahydrofolate Reductase	wlm96.pk047.l4	7	8

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Research 13*:3021-3030 (1985) and in the *Biochemical Journal 219 (No. 2)*:345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

DETAILED DESCRIPTION OF THE INVENTION

In the context of this disclosure, a number of terms shall be utilized. As used herein, a "nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. A nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

As used herein, "substantially similar" refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the polypeptide encoded by the nucleotide sequence. "Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by gene silencing through for example antisense or co-suppression technology. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate gene silencing or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention encompasses more than the specific exemplary nucleotide or amino acid sequences and includes functional equivalents thereof.

For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a nucleic acid fragment which result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded polypeptide, are well known in the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the polypeptide molecule would also not be expected to alter the activity of the polypeptide. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those

skilled in the art (Hames and Higgins, Eds. (1985) Nucleic Acid Hybridisation, IRL Press, Oxford, U.K.). Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2x SSC, 0.5% SDS was increased to 60°C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65°C.

Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent identity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by those skilled in this art. Preferred are those nucleic acid fragments whose nucleotide sequences encode amino acid sequences that are 85% identical to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are 90% identical to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are 95% identical to the amino acid sequences reported herein. Sequence alignments and percent identity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS. 5*:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

A "substantial portion" of an amino acid or nucleotide sequence comprises an amino acid or a nucleotide sequence that is sufficient to afford putative identification of the protein or gene that the amino acid or nucleotide sequence comprises. Amino acid and nucleotide sequences can be evaluated either manually by one skilled in the art, or by using computer-based sequence comparison and identification tools that employ algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol. 215*:403-410; see also www.ncbi.nlm.nih.gov/BLAST/). In general, a sequence of ten or more contiguous amino acids or thirty or more contiguous nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 30 or more contiguous nucleotides may be used in sequence-dependent methods

of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12 or more nucleotides may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises a nucleotide sequence that will afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches amino acid and nucleotide sequences encoding polypeptides that comprise one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

"Codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment comprising a nucleotide sequence that encodes all or a substantial portion of the amino acid sequences set forth herein. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid fragment for improved expression in a host cell, it is desirable to design the nucleic acid fragment such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic nucleic acid fragments" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form larger nucleic acid fragments which may then be enzymatically assembled to construct the entire desired nucleic acid fragment. "Chemically synthesized", as related to nucleic acid fragment, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of nucleic acid fragments may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the nucleic acid fragments can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers any gene that is not a native gene,

comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

"Coding sequence" refers to a nucleotide sequence that codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

"Promoter" refers to a nucleotide sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a nucleotide sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleotide segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. Promoters which cause a nucleic acid fragment to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg (1989) *Biochemistry of Plants 15*:1-82. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, nucleic acid fragments of different lengths may have identical promoter activity.

The "translation leader sequence" refers to a nucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA,

mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner and Foster (1995) *Molecular Biotechnology 3*:225).

The "3' non-coding sequences" refer to nucleotide sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al. (1989) *Plant Cell 1*:671-680.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into polypeptide by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to an RNA transcript that includes the mRNA and so can be translated into a polypeptide by the cell. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (see U.S. Patent No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific nucleotide sequence, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.

The term "operably linked" refers to the association of two or more nucleic acid fragments on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e.. that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference).

"Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

"Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

A "chloroplast transit peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast transit peptide. A "signal peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the secretory system (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol. 42*:21-53). If the protein is to be directed to a vacuole, a vacuolar targeting signal (*supra*) can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal (*supra*) may be added. If the protein is to be directed to the nucleus, any signal peptide present should be removed and instead a nuclear localization signal included (Raikhel (1992) *Plant Phys. 100*:1627-1632).

"Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol. 143*:277) and particle-accelerated or "gene gun" transformation technology (Klein et al. (1987) *Nature (London) 327*:70-73; U.S. Patent No. 4,945,050, incorporated herein by reference).

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Maniatis").

Nucleic acid fragments encoding at least a portion of a tetrahydrofolate metabolism enzyme have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. The nucleic acid fragments of the instant invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA

amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain reaction).

For example, genes encoding other 5,10-methylenetetrahydrofolate reductase enzymes, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al. (1988) *Proc. Natl. Acad. Sci. USA 85*:8998) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al. (1989) *Proc. Natl. Acad. Sci. USA 86*:5673; Loh et al. (1989) *Science 243*:217). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman and Martin (1989) *Techniques 1*:165).

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner (1984) *Adv. Immunol. 36*:1; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed polypeptides are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the level of folic acid in those cells.

Overexpression of the proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

Plasmid vectors comprising the instant chimeric gene can then constructed. The choice of plasmid vector is dependent upon the method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al. (1985) *EMBO J. 4*:2411-2418; De Almeida et al. (1989) *Mol. Gen. Genetics 218*:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

For some applications it may be useful to direct the instant polypeptides to different cellular compartments, or to facilitate its secretion from the cell. It is thus envisioned that the chimeric gene described above may be further supplemented by altering the coding sequence to encode the instant polypeptides with appropriate intracellular targeting sequences such as transit sequences (Keegstra (1989) *Cell 56*:247-253), signal sequences or sequences encoding endoplasmic reticulum localization (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol. 42*:21-53), or nuclear localization signals (Raikhel (1992) *Plant Phys. 100*:1627-1632) added and/or with targeting sequences that are already present removed. While the references cited give examples of each of these, the list is not exhaustive and more targeting signals of utility may be discovered in the future.

It may also be desirable to reduce or eliminate expression of genes encoding the instant polypeptides in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant polypeptide can be constructed by linking a gene or gene fragment encoding that polypeptide to plant promoter sequences. Alternatively, a chimeric gene designed to express antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in

reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

Molecular genetic solutions to the generation of plants with altered gene expression have a decided advantage over more traditional plant breeding approaches. Changes in plant phenotypes can be produced by specifically inhibiting expression of one or more genes by antisense inhibition or cosuppression (U. S. Patent Nos. 5,190,931, 5,107,065 and 5,283,323). An antisense or cosuppression construct would act as a dominant negative regulator of gene activity. While conventional mutations can yield negative regulation of gene activity these effects are most likely recessive. The dominant negative regulation available with a transgenic approach may be advantageous from a breeding perspective. In addition, the ability to restrict the expression of specific phenotype to the reproductive tissues of the plant by the use of tissue specific promoters may confer agronomic advantages relative to conventional mutations which may have an effect in all tissues in which a mutant gene is ordinarily expressed.

The person skilled in the art will know that special considerations are associated with the use of antisense or cosuppresion technologies in order to reduce expression of particular genes. For example, the proper level of expression of sense or antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan. Once transgenic plants are obtained by one of the methods described above, it will be necessary to screen individual transgenics for those that most effectively display the desired phenotype. Accordingly, the skilled artisan will develop methods for screening large numbers of transformants. The nature of these screens will generally be chosen on practical grounds, and is not an inherent part of the invention. For example, one can screen by looking for changes in gene expression by using antibodies specific for the protein encoded by the gene being suppressed, or one could establish assays that specifically measure enzyme activity. A preferred method will be one which allows large numbers of samples to be processed rapidly, since it will be expected that a large number of transformants will be negative for the desired phenotype.

The instant polypeptides (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to the these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting the polypeptides of the instant invention *in situ* in cells or *in vitro* in cell extracts. Preferred heterologous host cells for production of the instant polypeptides are microbial hosts. Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct a chimeric gene for production of the instant polypeptides. This chimeric gene could then be introduced into appropriate

microorganisms via transformation to provide high level expression of the encoded tetrahydrofolate metabolism enzyme. An example of a vector for high level expression of the instant polypeptides in a bacterial host is provided (Example 6).

Additionally, the instant polypeptides can be used as a targets to facilitate design and/or identification of inhibitors of those enzymes that may be useful as herbicides. This is desirable because the polypeptides described herein catalyze various steps in folic acid biosynthesis. Accordingly, inhibition of the activity of one or more of the enzymes described herein could lead to inhibition plant growth. Thus, the instant polypeptides could be appropriate for new herbicide discovery and design.

All or a substantial portion of the nucleic acid fragments of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al. (1987) *Genomics 1*:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted and used to calculate the position of the instant nucleic acid sequence in the genetic map previously obtained using this population (Botstein et al. (1980) *Am. J. Hum. Genet. 32*:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in Bernatzky and Tanksley (1986) *Plant Mol. Biol. Reporter 4(1)*:37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used for physical mapping (i.e., placement of sequences on physical maps; *see* Hoheisel et al. In: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, and references cited therein).

In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask (1991) *Trends Genet. 7*:149-154). Although current methods of FISH mapping favor use of

large clones (several to several hundred KB; see Laan et al. (1995) *Genome Research* *5*:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian (1989) *J. Lab. Clin. Med. 114(2)*:95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield et al. (1993) *Genomics* *16*:325-332), allele-specific ligation (Landegren et al. (1988) *Science 241*:1077-1080), nucleotide extension reactions (Sokolov (1990) *Nucleic Acid Res. 18*:3671), Radiation Hybrid Mapping (Walter et al. (1997) *Nature Genetics* 7:22-28) and Happy Mapping (Dear and Cook (1989) *Nucleic Acid Res. 17*:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these genes contained in a maize population carrying mutations in all possible genes (Ballinger and Benzer (1989) *Proc. Natl. Acad. Sci USA 86*:9402; Koes et al. (1995) *Proc. Natl. Acad. Sci USA 92*:8149; Bensen et al. (1995) *Plant Cell* 7:75). The latter approach may be accomplished in two ways. First, short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The amplification of a specific DNA fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the instant polypeptides. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as that for a restriction enzyme site-anchored synthetic adaptor. With either method, a plant containing a mutation in the endogenous gene encoding the instant polypeptides can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the instant polypeptides disclosed herein.

EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be

understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1

Composition of cDNA Libraries: Isolation and Sequencing of cDNA Clones

cDNA libraries representing mRNAs from various corn, rice, soybean and wheat tissues were prepared. The characteristics of the libraries are described below.

TABLE 2

cDNA Libraries from Corn, Rice, Soybean and Wheat

Library	Tissue	Clone
cco1	Corn (*Zea mays* L.) cob of 67 day old plants grown in green house	cco1.pk0049.d4
rls48	Rice (*Oryza sativa* L.) leaf (15 days after germination) 48 hours after infection of *Magaporthe grisea* strain 4360-R-67 (avr2-yamo); Susceptible	rls48.pk0005.b11
sfl1	Soybean (*Glycine max* L.) immature flower	sfl1.pk0017.d12
wlm96	Wheat (Triticum aestivum L.) seedlings 96 hr after inoculation w/ *E. graminis*	wlm96.pk047.l4

cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP™ XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut Bluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams et al., (1991) *Science 252*:1651). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

EXAMPLE 2

Identification of cDNA Clones

cDNA clones encoding tetrahydrofolate metabolism enzymes were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol. 215*:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States (1993) *Nature Genetics 3*:266-272) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

EXAMPLE 3

Characterization of cDNA Clones Encoding 5,10-Methylenetetrahydrofolate Reductase

The BLASTX search using the EST sequences from clones listed in Table 3 revealed similarity of the polypeptides encoded by the cDNAs to 5,10-methylenetetrahydrofolate reductase from *Schizosaccharomyces pombe* (NCBI Identifier No. gi 3702200) and *Homo sapiens* (NCBI Identifier Nos. gi 4336819 and gi 4753778). Shown in Table 3 are the BLAST results for individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), or contigs assembled from two or more ESTs ("Contig"):

TABLE 3

BLAST Results for Sequences Encoding Polypeptides Homologous to *Schizosaccharomyces pombe* and *Homo sapiens* 5,10-Methylenetetrahydrofolate Reductase

Clone	Status	BLAST pLog Score
cco1.pk0049.d4	FIS	26.70 (gi 3702200)
rls48.pk0005.b11	EST	104.00 (gi 4336819)
sfl1.pk0017.d12	EST	104.00 (gi 4336819)
wlm96.pk047.l4	EST	20.30 (gi 4753778)

The data in Table 4 represents a calculation of the percent identity of the amino acid sequences set forth in SEQ ID NOs:2, 4, 6 and 8 the *Schizosaccharomyces pombe* and *Homo sapiens* sequences.

TABLE 4

Percent Identity of Amino Acid Sequences Deduced From the Nucleotide Sequences of cDNA Clones Encoding Polypeptides Homologous to *Schizosaccharomyces pombe* and *Homo sapiens* 5,10-Methylenetetrahydrofolate Reductase

SEQ ID NO.	Percent Identity to
2	37% (gi 3702200)
4	41% (gi 4336819)
6	41% (gi 4336819)
8	51% (gi 4753778)

The instant amino acid sequences were also found to be similar to an amino acid sequence encoding an unknown protein from *Arabidopsis thaliana* (NCBI Identifier No. gi 3212869).

Sequence alignments and percent identity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS.* *5*:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant cDNA clones encode a substantial portion of a 5,10-methylenetetrahydrofolate reductase. These sequences represent the first plant sequences encoding 5,10-methylenetetrahydrofolate reductase.

EXAMPLE 4

Expression of Chimeric Genes in Monocot Cells

A chimeric gene comprising a cDNA encoding the instant polypeptides in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate

band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection. 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb SalI-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-SalI fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit; U.S. Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding the instant polypeptides, and the 10 kD zein 3' region.

The chimeric gene described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 μm in diameter) are coated with DNA using the following technique. Ten μg of plasmid DNAs are added to 50 μL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 μL of a 2.5 M solution) and spermidine free base (20 μL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After

10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 μL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 μL of ethanol. An aliquot (5 μL) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains gluphosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing gluphosinate. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology* *8*:833-839).

EXAMPLE 5

Expression of Chimeric Genes in Dicot Cells

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem. 261*:9228-9238) can be used for expression of the instant polypeptides in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

Soybean embroys may then be transformed with the expression vector comprising sequences encoding the instant polypeptides. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) *327*:70, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature 313*:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al.(1983) *Gene 25*:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the instant polypeptides and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 μL of a 60 mg/mL 1 μm gold particle suspension is added (in order): 5 μL DNA (1 μg/μL), 20 μl spermidine (0.1 M), and 50 μL $CaCl_2$ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 μL 70% ethanol and resuspended in 40 μL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five μL of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

EXAMPLE 6

Expression of Chimeric Genes in Microbial Cells

The cDNAs encoding the instant polypeptides can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987) *Gene 56*:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 μg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 μL of water. Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized

with phenol/chloroform as described above. The prepared vector pBT430 and fragment can then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 µg/mL ampicillin. Transformants containing the gene encoding the instant polypeptides are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol. 189*:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio-β-galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 µL of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One µg of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

EXAMPLE 7

Evaluating Compounds for Their Ability to Inhibit the Activity of Tetrahydrofolate Metabolism Enzymes

The polypeptides described herein may be produced using any number of methods known to those skilled in the art. Such methods include, but are not limited to, expression in bacteria as described in Example 6, or expression in eukaryotic cell culture, *in planta*, and using viral expression systems in suitably infected organisms or cell lines. The instant polypeptides may be expressed either as mature forms of the proteins as observed *in vivo* or as fusion proteins by covalent attachment to a variety of enzymes, proteins or affinity tags. Common fusion protein partners include glutathione S-transferase ("GST"), thioredoxin ("Trx"), maltose binding protein, and C- and/or N-terminal hexahistidine polypeptide ("$(His)_6$"). The fusion proteins may be engineered with a protease recognition site at the fusion point so that fusion partners can be separated by protease digestion to yield intact mature enzyme. Examples of such proteases include thrombin, enterokinase and factor Xa. However, any protease can be used which specifically cleaves the peptide connecting the fusion protein and the enzyme.

Purification of the instant polypeptides, if desired, may utilize any number of separation technologies familiar to those skilled in the art of protein purification. Examples of such methods include, but are not limited to, homogenization, filtration, centrifugation,

heat denaturation, ammonium sulfate precipitation, desalting, pH precipitation, ion exchange chromatography, hydrophobic interaction chromatography and affinity chromatography, wherein the affinity ligand represents a substrate, substrate analog or inhibitor. When the instant polypeptides are expressed as fusion proteins, the purification protocol may include the use of an affinity resin which is specific for the fusion protein tag attached to the expressed enzyme or an affinity resin containing ligands which are specific for the enzyme. For example, the instant polypeptides may be expressed as a fusion protein coupled to the C-terminus of thioredoxin. In addition, a $(His)_6$ peptide may be engineered into the N-terminus of the fused thioredoxin moiety to afford additional opportunities for affinity purification. Other suitable affinity resins could be synthesized by linking the appropriate ligands to any suitable resin such as Sepharose-4B. In an alternate embodiment, a thioredoxin fusion protein may be eluted using dithiothreitol; however, elution may be accomplished using other reagents which interact to displace the thioredoxin from the resin. These reagents include β–mercaptoethanol or other reduced thiol. The eluted fusion protein may be subjected to further purification by traditional means as stated above, if desired. Proteolytic cleavage of the thioredoxin fusion protein and the enzyme may be accomplished after the fusion protein is purified or while the protein is still bound to the ThioBond™ affinity resin or other resin.

Crude, partially purified or purified enzyme, either alone or as a fusion protein, may be utilized in assays for the evaluation of compounds for their ability to inhibit enzymatic activation of the instant polypeptides disclosed herein. Assays may be conducted under well known experimental conditions which permit optimal enzymatic activity. For example, assays for 5,10-methylenetetrahydrofolate reductase are presented by (West et al, (1993) *J. Biol. Chem. 268*:153-160 and D'Ari et al. (1991) *J. Biol. Chem. 266*:23953-23958).

CLAIMS

What is claimed is:

1. An isolated nucleic acid fragment encoding a 5,10-methylenetetrahydrofolate reductase comprising a member selected from the group consisting of:
   (a) an isolated nucleic acid fragment encoding an amino acid sequence that is at least 85% identical to the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 4, 6 and 8;
   (b) an isolated nucleic acid fragment that is complementary to (a).

2. The isolated nucleic acid fragment of Claim 1 wherein nucleic acid fragment is a functional RNA.

3. The isolated nucleic acid fragment of Claim 1 wherein the nucleotide sequence of the fragment comprises the sequence set forth in a member selected from the group consisting of SEQ ID NO:1, 3, 5 and 7.

4. A chimeric gene comprising the nucleic acid fragment of Claim 1 operably linked to suitable regulatory sequences.

5. A transformed host cell comprising the chimeric gene of Claim 4.

6. A 5,10-methylenetetrahydrofolate reductase polypeptide comprising all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, 4, 6 and 8.

7. A method of altering the level of expression of a tetrahydrofolate metabolism enzyme in a host cell comprising:
   (a) transforming a host cell with the chimeric gene of Claim 4; and
   (b) growing the transformed host cell produced in step (a) under conditions that are suitable for expression of the chimeric gene

wherein expression of the chimeric gene results in production of altered levels of a tetrahydrofolate metabolism enzyme in the transformed host cell.

8. A method of obtaining a nucleic acid fragment encoding all or a substantial portion of the amino acid sequence encoding a tetrahydrofolate metabolism enzyme comprising:
   (a) probing a cDNA or genomic library with the nucleic acid fragment of Claim 1;
   (b) identifying a DNA clone that hybridizes with the nucleic acid fragment of Claim 1;
   (c) isolating the DNA clone identified in step (b); and
   (d) sequencing the cDNA or genomic fragment that comprises the clone isolated in step (c)

wherein the sequenced nucleic acid fragment encodes all or a substantial portion of the amino acid sequence encoding a tetrahydrofolate metabolism enzyme.

9. A method of obtaining a nucleic acid fragment encoding a substantial portion of an amino acid sequence encoding a tetrahydrofolate metabolism enzyme comprising:

(a) synthesizing an oligonucleotide primer corresponding to a portion of the sequence set forth in any of SEQ ID NOs:1, 3, 5 and 7; and

(b) amplifying a cDNA insert present in a cloning vector using the oligonucleotide primer of step (a) and a primer representing sequences of the cloning vector

wherein the amplified nucleic acid fragment encodes a substantial portion of an amino acid sequence encoding a tetrahydrofolate metabolism enzyme.

10. The product of the method of Claim 8.

11. The product of the method of Claim 9.

12. A method for evaluating at least one compound for its ability to inhibit the activity of a tetrahydrofolate metabolism enzyme, the method comprising the steps of:

(a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a tetrahydrofolate metabolism enzyme, operably linked to suitable regulatory sequences;

(b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of the tetrahydrofolate metabolism enzyme encoded by the operably linked nucleic acid fragment in the transformed host cell;

(c) optionally purifying the tetrahydrofolate metabolism enzyme expressed by the transformed host cell;

(d) treating the tetrahydrofolate metabolism enzyme with a compound to be tested; and

(e) comparing the activity of the tetrahydrofolate metabolism enzyme that has been treated with a test compound to the activity of an untreated tetrahydrofolate metabolism enzyme,

thereby selecting compounds with potential for inhibitory activity.

SEQUENCE LISTING

<110> Falco, Carl
Famodu, Layo O.

<120> Tetrahydrofolate Metabolism Enzymes

<130> BB-1179

<140>
<141>

<150> 60/092,869
<151> July 15, 1998

<160> 8

<170> Microsoft Office 97

<210> 1
<211> 542
<212> DNA
<213> Zea mays

<220>
<221> unsure
<222> (363)

<400> 1
```
gcacgagcat tagagcgcgc gcgccaagct cgtggtgagg ttttcgagat gaaggttatc  60
gagaagatcc tggaggcggc gggggatggc cggacggcgt tctcgttcga gtacttccct 120
cccaagacgg aggagggggt cgagaacctg ttcgagcgga tggaccgcat ggtggcgcac 180
ggcccctcct tctgcgacat cacctggggc gccgggggat ccaccgctga ccttaccctc 240
gaaatcgcca accgtatgca gaacatggtg tgtgtggaaa ccatgatgca cctgacatgt 300
acaaacatgc cagtagagaa gattgaccat gccttggaaa ccatcaaatc caatgggatt 360
canaatgttt tggccctcag aggggattct ccacatgggc aagacaaatt tgtgcaagtt 420
gaaggcggat ttgcttgtgc tcttgatttg gtgacacata ttatagccaa gtacggtgaa 480
ttattttggc atactgttac tggttatcag aagcacacct tgagcgatac atgcgaggga 540
gg                                                                542
```

<210> 2
<211> 164
<212> PRT
<213> Zea mays

<220>
<221> UNSURE
<222> (105)

<400> 2
```
Met Lys Val Ile Glu Lys Ile Leu Glu Ala Ala Gly Asp Gly Arg Thr
 1               5                  10                  15

Ala Phe Ser Phe Glu Tyr Phe Pro Pro Lys Thr Glu Glu Gly Val Glu
            20                  25                  30

Asn Leu Phe Glu Arg Met Asp Arg Met Val Ala His Gly Pro Ser Phe
        35                  40                  45
```

```
Cys Asp Ile Thr Trp Gly Ala Gly Gly Ser Thr Ala Asp Leu Thr Leu
    50                  55                  60

Glu Ile Ala Asn Arg Met Gln Asn Met Val Cys Val Glu Thr Met Met
65                  70                  75                  80

His Leu Thr Cys Thr Asn Met Pro Val Glu Lys Ile Asp His Ala Leu
                85                  90                  95

Glu Thr Ile Lys Ser Asn Gly Ile Xaa Asn Val Leu Ala Leu Arg Gly
            100                 105                 110

Asp Ser Pro His Gly Gln Asp Lys Phe Val Gln Val Glu Gly Gly Phe
        115                 120                 125

Ala Cys Ala Leu Asp Leu Val Thr His Ile Ile Ala Lys Tyr Gly Glu
    130                 135                 140

Leu Phe Trp His Thr Val Thr Gly Tyr Gln Lys His Thr Leu Ser Asp
145                 150                 155                 160

Thr Cys Glu Gly


<210>  3
<211>  591
<212>  DNA
<213>  Oryza sativa

<220>
<221>  unsure
<222>  (535)

<220>
<221>  unsure
<222>  (548)

<220>
<221>  unsure
<222>  (554)

<220>
<221>  unsure
<222>  (557)

<220>
<221>  unsure
<222>  (561)

<220>
<221>  unsure
<222>  (591)

<400>  3
tctaccttac cttacctcac ctcacaaggg acaaagataa aaaatgaaga tcattgagaa  60
aattcacgca gcctcggctg atccaaacag ggtggtgttc tcatttgagt tcttcccgcc 120
caagactgag gatggggtgg acaatctgtt tgagaggatg gaccgcatgg tggtgcacaa 180
cccctcgttt tgcgacatca cttggggtgc tggagggaca acggctgatc ttacattgga 240
aattgccaac aagatgcaga acattgtctg tgtggagacc atgatgcacc tcacctgcac 300
caacatgcct gttgagaaga ttgaccatgc tctccacacc atcaagtcca atggcctcca 360
```

aaatgtgctt gcacttcgag gtgatccacc gcatggtcag gacaaattcg tccagtcgaa 420
ggcggtttcc tgtgcacgcg acttggtgca acatatcaga ctaaatatgg tgacacttgg 480
ataagtcctg gtatcagagc acaccggatg tatagaatga tggttggcac acagnggtac 540
aaatatcnct actnaanaca ngttatctgg acattcgatg caccattttt n 591

<210> 4
<211> 408
<212> PRT
<213> Oryza sativa

<400> 4
Ala Arg Val Ile Val Thr Gln Leu Phe Tyr Asp Thr Asp Ile Phe Leu
1 5 10 15

Lys Phe Val Asn Asp Cys Arg Gln Ile Gly Ile Thr Cys Pro Ile Val
20 25 30

Pro Gly Ile Met Pro Ile Asn Asn Tyr Lys Gly Phe Ile Arg Met Thr
35 40 45

Gly Phe Cys Lys Thr Lys Ile Pro Ala Asp Ile Met Ala Ala Leu Glu
50 55 60

Pro Ile Lys Asp Asn Glu Glu Ala Val Lys Ala Tyr Gly Ile His Leu
65 70 75 80

Gly Thr Glu Met Cys Lys Lys Ile Leu Ala His Gly Ile Lys Thr Leu
85 90 95

His Leu Tyr Thr Leu Asn Met Glu Lys Ser Ala Leu Ala Ile Leu Met
100 105 110

Asn Leu Gly Leu Ile Glu Glu Ser Lys Val Ser Arg Ser Leu Pro Trp
115 120 125

Arg Arg Pro Ala Asn Val Phe Arg Val Lys Glu Asp Val Arg Pro Ile
130 135 140

Phe Trp Ala Asn Arg Pro Lys Ser Tyr Ile Ser Arg Thr Ile Gly Trp
145 150 155 160

Asp Gln Tyr Pro His Gly Arg Trp Gly Asp Ser Cys Asn Pro Ser Tyr
165 170 175

Gly Ala Leu Ser Asp Tyr Gln Phe Met Arg Pro Arg Ala Arg Asp Lys
180 185 190

Lys Leu Val Glu Glu Trp Ala Val Pro Leu Lys Ser Val Glu Asp Ile
195 200 205

Tyr Glu Arg Phe Arg Leu Tyr Cys Leu Gly Lys Leu Arg Ser Asn Pro
210 215 220

Trp Ser Glu Leu Asp Gly Leu Gln Pro Glu Thr Lys Ile Ile Asn Glu
225 230 235 240

Gln Leu Glu Lys Ile Asn Thr Lys Gly Phe Leu Thr Ile Asn Ser Gln
245 250 255

Pro Ala Val Asn Gly Glu Lys Ser Asp Ser Pro Thr Val Gly Trp Gly
260 265 270

Gly Pro Gly Gly Tyr Val Tyr Gln Lys Ala Tyr Val Glu Phe Phe Cys
275 280 285

Ser Lys Glu Lys Leu Asp Ala Leu Val Asp Lys Cys Lys Asp Arg Thr
290 295 300

Ser Leu Thr Tyr Met Ala Val Asn Lys Asp Gly Ser Trp Lys Ser Asn
305 310 315 320

Val Gly Gln Thr Asp Val Asn Ala Val Thr Trp Gly Val Phe Pro Ala
325 330 335

Lys Glu Ile Ile Gln Pro Thr Ile Val Asp Pro Val Ser Phe Asn Val
340 345 350

Trp Lys Asp Glu Ala Phe Glu Ile Trp Ser Arg Gly Trp Ala Ser Leu
355 360 365

Tyr Pro Glu Asp Glu Ala Ser Arg Lys Leu Val Glu Glu Val Gly Gly
370 375 380

Ser His Phe Leu Val Ser Leu Val Asp Asn Asp Tyr Ile Asn Gly Asp
385 390 395 400

Leu Phe Ala Val Phe Ala Asp Phe
405

<210> 5
<211> 558
<212> DNA
<213> Glycine max

<400> 5
tgattgtcac ccaattattt tatgatacgg atatattcct caaatttgtg aacgactgtc 60
gccaaattgg aataacgtgt cctattgtac ctggaattat gcccattaat aattacaagg 120
gctttatccg catgactggg ttttgcaaaa caaagatacc agctgacatt atggctgctt 180
tagagcctat caaggacaat gaagaagctg tcaaggctta tggaattcac ctgggaactg 240
aaatgtgcaa aaagatttta gctcatggaa ttaagacatt gcatctttat acactaaata 300
tggagaaatc tgcattggca atactaatga accttggcct aattgaagag tccaaagttt 360
ctaggtcctt accttggaga cgccctgcaa atgttttccg tgttaaagaa gatgtccgtc 420
caatcttttg ggcaaatcga ccaaaaagct acatatcaag gaccatagga tgggatcaat 480
acccacatgg gcgttggggt gattcctgta atccatcata tggtgcatta tctgattatc 540
agttcatgcg gccacgtg 558

<210> 6
<211> 408
<212> PRT
<213> Glycine max

<400> 6
Ala Arg Val Ile Val Thr Gln Leu Phe Tyr Asp Thr Asp Ile Phe Leu
1 5 10 15

Lys Phe Val Asn Asp Cys Arg Gln Ile Gly Ile Thr Cys Pro Ile Val
20 25 30

Pro Gly Ile Met Pro Ile Asn Asn Tyr Lys Gly Phe Ile Arg Met Thr
35 40 45

Gly Phe Cys Lys Thr Lys Ile Pro Ala Asp Ile Met Ala Ala Leu Glu
50 55 60

Pro Ile Lys Asp Asn Glu Glu Ala Val Lys Ala Tyr Gly Ile His Leu
65 70 75 80

Gly Thr Glu Met Cys Lys Lys Ile Leu Ala His Gly Ile Lys Thr Leu
85 90 95

His Leu Tyr Thr Leu Asn Met Glu Lys Ser Ala Leu Ala Ile Leu Met
100 105 110

Asn Leu Gly Leu Ile Glu Glu Ser Lys Val Ser Arg Ser Leu Pro Trp
115 120 125

Arg Arg Pro Ala Asn Val Phe Arg Val Lys Glu Asp Val Arg Pro Ile
130 135 140

Phe Trp Ala Asn Arg Pro Lys Ser Tyr Ile Ser Arg Thr Ile Gly Trp
145 150 155 160

Asp Gln Tyr Pro His Gly Arg Trp Gly Asp Ser Cys Asn Pro Ser Tyr
165 170 175

Gly Ala Leu Ser Asp Tyr Gln Phe Met Arg Pro Arg Ala Arg Asp Lys
180 185 190

Lys Leu Val Glu Glu Trp Ala Val Pro Leu Lys Ser Val Glu Asp Ile
195 200 205

Tyr Glu Arg Phe Arg Leu Tyr Cys Leu Gly Lys Leu Arg Ser Asn Pro
210 215 220

Trp Ser Glu Leu Asp Gly Leu Gln Pro Glu Thr Lys Ile Ile Asn Glu
225 230 235 240

Gln Leu Glu Lys Ile Asn Thr Lys Gly Phe Leu Thr Ile Asn Ser Gln
245 250 255

Pro Ala Val Asn Gly Glu Lys Ser Asp Ser Pro Thr Val Gly Trp Gly
260 265 270

Gly Pro Gly Gly Tyr Val Tyr Gln Lys Ala Tyr Val Glu Phe Phe Cys
275 280 285

Ser Lys Glu Lys Leu Asp Ala Leu Val Asp Lys Cys Lys Asp Arg Thr
290 295 300

Ser Leu Thr Tyr Met Ala Val Asn Lys Asp Gly Ser Trp Lys Ser Asn
305 310 315 320

Val Gly Gln Thr Asp Val Asn Ala Val Thr Trp Gly Val Phe Pro Ala
325 330 335

Lys Glu Ile Ile Gln Pro Thr Ile Val Asp Pro Val Ser Phe Asn Val
340 345 350

Trp Lys Asp Glu Ala Phe Glu Ile Trp Ser Arg Gly Trp Ala Ser Leu
355 360 365

Tyr Pro Glu Asp Glu Ala Ser Arg Lys Leu Val Glu Glu Val Gly Gly
370 375 380

Ser His Phe Leu Val Ser Leu Val Asp Asn Asp Tyr Ile Asn Gly Asp
385 390 395 400

Leu Phe Ala Val Phe Ala Asp Phe
405

<210> 7
<211> 451
<212> DNA
<213> Triticum aestivum

<220>
<221> unsure
<222> (25)

<220>
<221> unsure
<222> (115)

<220>
<221> unsure
<222> (120)

<220>
<221> unsure
<222> (176)

<220>
<221> unsure
<222> (347)

<220>
<221> unsure
<222> (371)

<220>
<221> unsure
<222> (426)

<220>
<221> unsure
<222> (442)

<220>
<221> unsure
<222> (445)

<400> 7
ggagcacatt aaagccaagt atggngatta ctttggcata actgtcgctg gctatccaga 60
ggcacaccct gaggtaatac taacgcgagg aaggtgctac ggaggaagca tatancaaan 120
atcttgctta cttgaagaga aaggttgatg ctggtgctga cgttataatc acccancttt 180
tctatgatac cgatatcttt ctcaagtttg tgaacgactg ccgtcagatt ggtataacct 240
gccctatcgt tcctggcata atgccaataa ataactacaa aggatttgtg cgcatgactg 300
gattctgcaa aactaaaatt ccacctgaga ttactgctgc cttgggntcc tactaaagac 360

aatgaggagg ntgtgaaaag catatgggat ccaccctggt actgaagatg ttcaaaaaa 420
attttnggct agtgggataa anacnttgca c 451

<210> 8
<211> 101
<212> PRT
<213> Triticum aestivum

<220>
<221> UNSURE
<222> (25)..(26)

<220>
<221> UNSURE
<222> (29)

<220>
<221> UNSURE
<222> (31)

<220>
<221> UNSURE
<222> (49)

<400> 8
Glu His Ile Lys Ala Lys Tyr Gly Asp Tyr Phe Gly Ile Thr Val Ala
1 5 10 15

Gly Tyr Pro Glu Ala His Pro Glu Xaa Xaa Ala Tyr Xaa Lys Xaa Leu
20 25 30

Ala Tyr Leu Lys Arg Lys Val Asp Ala Gly Ala Asp Val Ile Ile Thr
35 40 45

Xaa Leu Phe Tyr Asp Thr Asp Ile Phe Leu Lys Phe Val Asn Asp Cys
50 55 60

Arg Gln Ile Gly Ile Thr Cys Pro Ile Val Pro Gly Ile Met Pro Ile
65 70 75 80

Asn Asn Tyr Lys Gly Phe Val Arg Met Thr Gly Phe Cys Lys Thr Lys
85 90 95

Ile Pro Pro Glu Ile
100

INTERNATIONAL SEARCH REPORT

International Application No
~~PCT/US 99/15916~~

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/53 C12N9/06 C12N5/10 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SASAKI T.: "AC C26645" EMBL DATABASE,6 August 1997 (1997-08-06), XP002119521 Heidelberg the whole document ---	1,2,4-6
X	SASAKI T. ET AL.: "AC D48925" EMBL DATABASE,9 March 1995 (1995-03-09), XP002119522 Heidelberg the whole document ---	1,2,4-6
X	SASAKI T. ET AL.: "AC D46316" EMBL DATABASE,9 March 1995 (1995-03-09), XP002119523 Heidelberg the whole document --- -/--	1,2,4-6

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 October 1999	03/11/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Kania, T

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	NOUR J M ET AL: "Isolation and sequencing of the cDNA coding for spinach 10-formyltetrahydrofolate synthetase. Comparisons with the yeast, mammalian, and bacterial proteins." JOURNAL OF BIOLOGICAL CHEMISTRY, (1992 AUG 15) 267 (23) 16292-6. , XP002119524 the whole document ---	12
Y	PETERSON, CLARENCE DENIS: "Antifolate studies: the evaluation of folate analogs as inhibitors of formyltetrahydrofolate synthetase, methenyltetrahydrofolate cyclohydrolase and methylenetetrahydrofolate dehydrogenase" (1979) 109 PP. AVAIL.: UNIV. MICROFILMS INT., ORDER NO. 8000085 FROM: DISS. ABSTR. INT. B 1980, 40(9), 4278, XP002119525 abstract ---	12
A	BENNER, M. ET AL: "Amplification of the methylenetetrahydrofolate reductase-- gene" MAIZE GENETICS COOPERATION NEWSLETTER, (1995) NO. 69, PP. 89., 'Online! XP002119526 Science and Technology Centre, Rider University, Lawrenceville, New Jerse 08648-3099, USA. Retrieved from the Internet: <URL:www.agron.missouri.edu/mnl/69/101benner.html> 'retrieved on 1999-10-15! abstract ---	1-12
A	WO 95 33054 A (UNIV MCGILL ;ROZEN RIMA (CA); GOYETTE PHILIPPE (CA)) 7 December 1995 (1995-12-07) the whole document ---	1-12
A	COVITZ P. ET AL.: "AC AA660667" EMBL DATABASE, 14 November 1997 (1997-11-14), XP002119527 Heidelberg the whole document -----	1-5

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/15916

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9533054 A	07-12-1995	AU 2519895 A	21-12-1995
		CA 2191220 A	07-12-1995
		EP 0755450 A	29-01-1997